

---

## New lexical blends in *The Simpsons*: a formal analysis of English nonce formations and their French translations

Adam Renwick and Vincent Renner

---

**Electronic version**

URL: <http://journals.openedition.org/lexis/3829>

DOI: 10.4000/lexis.3829

ISSN: 1951-6215

**Publisher**

Université Jean Moulin - Lyon 3

**Electronic reference**

Adam Renwick and Vincent Renner, « New lexical blends in *The Simpsons*: a formal analysis of English nonce formations and their French translations », *Lexis* [Online], 14 | 2019, Online since 16 December 2019, connection on 20 April 2020. URL : <http://journals.openedition.org/lexis/3829> ; DOI : <https://doi.org/10.4000/lexis.3829>

---

This text was automatically generated on 20 April 2020.



Lexis is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

---

# New lexical blends in *The Simpsons*: a formal analysis of English nonce formations and their French translations

Adam Renwick and Vincent Renner

---

## Introduction

- 1 New words are not always coined with the purpose of becoming institutionalized, i.e. of becoming part of the lexicon of a community of speakers. As Hohenhaus [2005: 365] puts it, “*nonce* can be the first stage in a longer life-span of a word but need not be – and mostly it is also the last stage”. New words serve a variety of communicative purposes and a cline of likelihood of institutionalization may be posited, where the playful creation in (1) could be placed at one extreme (institutionalization is relatively very unlikely) and the terminological creation in (2) at the other extreme (institutionalization is relatively very likely):

(1) *purr* + *perplexing* > *purrplexing* (*The Simpsons*, Season 26, Episode 18)

(2) *adsorbed* + *atom* > *adatom* (physical chemistry)<sup>1</sup>

- 2 Studying lexical blends should be especially enlightening in this regard as a correlation between functional type and formal variation may be postulated. Our hypothesis is that playful nonce formations might be noticeably different in their formal characteristics than institutionalized blends as a class because the identification of the source elements and the construal of meaning can only take place online, i.e. during the actual perception of speech. By definition, the various source elements of a blend do not all appear in full and the operation of blending leads to various degrees of formal opacification. Haplogenic blends, which are coined through segment overlap, as illustrated in (3), can be considered to be minimally opaque while outputs resulting

from the clipping of polysyllabic elements to mono- or infra-syllabic constituents, as illustrated in (4), are maximally opaque:

- (3a) *hip-hop* + *opera* > *hip-hopera*
- (3b) *winter* + *interim* > *winterim*
- (4a) *columbite* + *tantalite* > *coltan*
- (4b) *binary* + *digit* > *bit*

- 3 Playful nonce formations are thus expected to be characterized by a lesser degree of formal opacity than institutionalized items. This would be in line with previous research by Ronneberger-Sibold [2006], who reports in her study of the relative transparency of blend types in German that 66% of the 612 units in her dataset of literary and journalistic nonce formations, but only 16% of the 220 units in her dataset of brand names, are maximally transparent, i.e. coined through haplologic blending or the full (phonemic) overlap of source elements (as in *Jewbilee*, from *Jew* and *jubilee*).
- 4 In order to assemble a large dataset of playful nonce blends in present-day English, we resorted to collecting items from a tailor-made corpus comprising 29 seasons of scripts from the US animated TV series *The Simpsons*, a long-running sitcom identified for its lexical creativity and expected to contain a sizable number of blends. As directly comparable data were available in Hexagonal French in the form of translated script, a lexical analysis of the translations of the English nonce blends into French was also subsequently carried out. Analytic data on English and French institutionalized blends taken from Renner [2019] were used as points of comparison.

## 1. Methodology

### 1.1. English data

- 5 To collect the novel lexical blends appearing in *The Simpsons*, we first compiled an electronic corpus of transcripts of the 639 episodes from the 29 complete seasons broadcast up to mid-2018. This corpus was based on fan-created transcripts published to a website dedicated to transcripts of television shows, <http://transcripts.foreverdreaming.org/>. It is of note that these transcripts do not necessarily exactly reflect the scripts used by the cast of the TV show when recording dialogue for the episodes and are subject to some misinterpretations as well as transcription and typographical errors. However, the consultation of a second website containing transcripts of the series (<https://www.springfieldspringfield.co.uk/>) and the examination of specific clips from the show available on various video sharing platforms guarantee that the corrected transcripts that were compiled are highly accurate records of the lexical content of the show. HTML coding was stripped to give raw text files for each season which were analyzed with the MonoConc Pro concordancer. These 29 files comprised some 1.5 million words (approximately 2,300 words per episode on average).
- 6 Given our focus on novel blends, we chose to proceed by the use of an exclusion corpus to remove already attested words. To create the exclusion corpus, we combined several existing stoplists of English words in order to limit the number of candidate forms. The first two stoplists were based on *Webster's Second International Dictionary* in the form of wordlists named *web2* (235,887 items) and *web2a* (76,206 items) which are included in

Unix/Linux-based operating systems. They encompass the large percentage of the English lexicon that has remained formally unchanged since the publication of this dictionary in 1934. They were supplemented by three further stoplists that include lexical innovations up to the first decade of the twenty-first century: the list of the 20,000 most common words in the English GoogleBooks corpus (<https://github.com/first20hours/google-10000-english/blob/master/20k.txt>), the list of the American English spelling variants of the 10,000 most common words in the corpus (<https://github.com/first20hours/google-10000-english/blob/master/google-10000-english-usa.txt>) and a customized list combining the wordlists developed by the SCOWL (*And Friends*) project (<http://wordlist.aspell.net/>). These five stoplists were merged to form a final exclusion corpus containing 402,933 words. Using our concordancer, we generated a list of all items used in the 29 seasons of *The Simpsons* that were not present in the exclusion corpus. This produced a set of 19,709 candidate forms, including 14,812 hapaxes.

- 7 These remaining candidates were then examined individually and discarded if they were simplex words or had been constructed through other word-formation processes like affixation or compounding. The boundaries of the concept of lexical blending have been subject to debate in the morphological literature (see e.g. López-Rúa [2004], Bauer [2012: 19-21], Beliaeva [2014: 45-47] and Renner [2015: 99-105] for a discussion). To make valid comparisons with institutionalized blends, we chose to adopt the definition of blending provided by Renner [2019: 29], namely that a blend is a constructed word which does not contain all of its source elements in full while satisfying the following three requirements: firstly, it does not contain a recurring word fragment which could be assimilated to a combining form (as in *slumpflation*, which can be considered a compound of *slump* and *-flation*, or *Margealicious*, a Simpsonian coinage that was analyzed as the compounding of *Marge* and *-licious* with the help of a linking vowel); secondly, it does not manifest single external shortening (as in the case of *blog*, which can be considered a clipped form of *weblog* rather than a blend); and, thirdly, it is not coined through the clipping of the initials of a majority of its source elements, as is the case of *zineb* (< *zinc* + *ethylene* + *bisdithiocarbamate*), because, from a prototype-based perspective, such a formation is closer to the class of initialisms than to that of blends. This final sifting stage led to the identification of a total of 237 blends in a corpus of about 1.5 million words, at an average of 158 blends occurring per million words.

## 1.2. French data

- 8 To examine the manner in which the English blends were rendered in French, we created a corpus based on the Hexagonal French version of *The Simpsons* along the same lines as for the English corpus. We gathered the existing transcripts from fan-submitted transcripts of the first 26 seasons ([www.simpsonspark.com/scripts.php](http://www.simpsonspark.com/scripts.php)), as those episodes in Seasons 27-29 had not yet been broadcast in Hexagonal French, meaning we are only able to offer a translational perspective on 217 novel English blends amongst the 237 analyzed. From the French transcripts, we then extracted the utterances corresponding to the 217 English blends. In cases where it was uncertain based from the transcript alone that the utterance in question contained a blend, we viewed the relevant part of the episode in question to determine the accuracy of the transcription as compared to the audio and to gather any further non-verbal information present onscreen before, during and following the utterance in question.

This method sufficed to determine whether a blend was indeed present, as well as the source elements used to form it, and it was also used in the few cases where no transcript was yet online for certain episodes in Season 26.

## 2. A formal analysis of English nonce blends

- 9 The dataset of English nonce blends comprises 237 items, including 4 three-element items (e.g. *Frightmarestein* (< *fright* + *nightmare* + *Frankenstein*)) which are not analyzed further in this section because of their marginal status. All blends are given in their orthographic form for simplicity's sake, but they were retrieved from spoken data and the formal analysis below is thus solely based on the phonological form of each item.

### 2.1. Lexical shortening

- 10 The distribution of the various types of lexical shortening for the 233 two-element blends of the *Simpsons* dataset is presented in Table 1 and contrasted with that available for institutionalized blends (Renner [2019: 33]).

Table 1: Distribution of English blends according to the type of lexical shortening

Type of lexical shortening	Institutionalized blends	Nonce blends
left-hand-side inner shortening (e.g. <i>Viagrogaine</i> < <i>Viagr[a]</i> + <i>Rogaine</i> )	24%	17% (40 items)
right-hand-side inner shortening (e.g. <i>fudgesicle</i> < <i>fudge</i> + [ <i>pop</i> ] <i>sicle</i> )	21%	55.5% (129 items)
double inner shortening (e.g. <i>lupper</i> < <i>lu[nch]</i> + [ <i>s</i> ] <i>upper</i> )	31%	11% (26 items)
double right-shortening (e.g. <i>cyborg</i> < <i>cyb[ernetic]</i> + <i>org[anism]</i> )	14%	–
double left-shortening (e.g. <i>cueabunga</i> < [ <i>barbe</i> ] <i>cue</i> + [ <i>cowa</i> ] <i>bunga</i> )	–	0.5% (1 item)
haplologic blending (e.g. <i>galgebra</i> < <i>g[al]</i> + <i>al[gebra]</i> )	7%	12.5% (29 items)
sandwich blending (e.g. <i>Rastafrogian</i> < <i>Rastaf[ar]ian</i> + <i>frog</i> )	1%	3.5% (8 items)
other (e.g. <i>ziram</i> < <i>zi[nc]</i> + [ <i>ca</i> ] <i>r</i> [ <i>b</i> ] <i>am[ate]</i> )	2%	–

- 11 The first striking difference between the two distributions is the absence of double right-shortened units like *cyborg* (< *cybernetic* + *organism*) or *perfin* (< *perforated* + *initial*)

in the set of nonce blends, highlighting that such items form a very atypical class of blends and are sometimes not even classified as blends, but as “clipping compounds” or “complex clippings”, because of their distinctive formal features (see e.g. Beliaeva [2019, Section 3.2.] and Renner [2019: 44] for a discussion). The dispreference for the pattern of double right-shortening in nonce blends may be explained in terms of (non-)recognizability of the source elements, each of them being canonically clipped to a monosyllable, which often causes the syllabic contour of either source element to be lost.

- 12 The other four non-marginal types of lexical shortening are found in the two sets of blends, but in a starkly different frequency order. The most notable difference lies in the overwhelming preference among nonce blends for the pattern of right-hand-side inner shortening, as illustrated in *fudgesicle* (< *fudge* + *popsicle*) and *mathnasium* (< *math* + *gymnasium*), which accounts for more than half of all items (55.5%). Correlatively, the proportion of double inner-shortened blends like *lupper* (< *lunch* + *supper*) and *Purgatraz* (< *purgatory* + *Alcatraz*) is dramatically lower in the nonce-blend set (11%). This underscores that the preference for the pattern of double inner shortening may not be universal, as is commonly believed (Beliaeva [2019, Section 3.3.]):

One obvious regularity that is postulated in the literature as a defining feature of blends is that most blends combine the initial part of one word with the final part of another.

- 13 This relative dispreference for double inner shortening may, here too, be explained in terms of relative recognizability: an output which has retained one source element in full is formally more transparent than an output made of two fragments. This explanation may also be used to account for the relatively high percentage of haplologic blends like *galgebra* (< *gal* + *algebra*) and *Gaybraham* (< *gay* + *Abraham*) in the set of nonce blends. As pointed out in the Introduction, the blends in this class are minimally opaque, i.e. display the highest possible degree of recognizability.

## 2.2. Segment overlap

- 14 Segment overlap is not a defining, but a prototypical feature of lexical blending. It typically appears word-medially, around the point of splicing of the source elements or word fragments, as illustrated by the nonce blends in (5):

- (5a) *bland* + *tandoori* > *bl[and + and]oori* > *blandoori*  
 (5b) *Lamborghini* + *Bugatti* > *Lambor[g + g]atti* > *Lamborgatti*

- 15 Identical segments may also appear non-medially in the two source elements, as in the Simpsonian blends *peachza* (< *peach* + *pizza*) and *forfty* (< *forty* + *fifty*), but this is not considered a case of overlap *sensu stricto* (see Renner [2019: 33] for a discussion). Table 2 provides the breakdown of the different lengths of medial phonemic overlap in the *Simpsons* dataset.

**Table 2: Distribution of English blends according to the length of medial phonemic overlap**

Number of overlapping segments	Number of nonce blends
0 (e.g. <i>grabulous</i> < <i>great</i> + <i>fabulous</i> )	88 (38%)

1 (e.g. /s/ in <i>croissandwich</i> < <i>croissant</i> + <i>sandwich</i> )	65 (28%)
2 (e.g. /ju:/ in <i>youbicle</i> < <i>you</i> + <i>cubicle</i> )	55 (23.5%)
3 (e.g. /meɪl/ in <i>femailman</i> < <i>female</i> + <i>mailman</i> )	17 (7.5%)
4 (e.g. /raɪk/ in <i>Prankenstein</i> < <i>prank</i> + <i>Frankenstein</i> )	6 (2.5%)
5 (e.g. /dekst/ in <i>poindextrose</i> < <i>poindexter</i> + <i>dextrose</i> )	2 (< 1%)

- 16 Overlap has been measured to be present in almost half of all blends (44%) in the case of institutionalized items (Renner [2019: 34]). Unsurprisingly, the proportion is higher in the case of nonce blends – 62% – as segment overlap serves to maximize the recognizability of the source elements.

### 2.3. Phonological split points

- 17 The act of clipping source elements before splicing the remaining fragments into a lexical blend leads to five different possible types of phonological split point, as illustrated in (6):<sup>2</sup>

(6a) at a syllable boundary: *ce.leb.ri.[ty] + fawn.ing* > *ce.leb.ri.fawn.ing*

(6b) at an onset-nucleus boundary: *Frink + [ma.n]i.ac* > *Frin.ki.ac*

(6c) at a nucleus-coda boundary: *de.tec.ti[ve] + pals* > *de.tec.ti.pals*

(6d) inside a complex onset: *smock + [a.p]ron* > *smock.ron*

(6e) inside a complex coda: *co.bal[t] + [vi.t]a.min* > *co.bal.a.min* (Renner [2019: 35])<sup>3</sup>

- 18 Table 3 shows that the distribution of the five split points among nonce blends is more marked than that of institutionalized blends (Renner [2019: 35]), with two-thirds of all non-overlap blends respecting syllable boundaries.

Table 3: Distribution of split points in English non-overlap blends

Location of split points	Institutionalized blends	Nonce blends
Syllable boundary	51%	65.5% (67 items)
Onset-nucleus boundary	35%	30.5% (31 items)
Nucleus-coda boundary	9.5%	3% (3 items)
Inside a complex onset	2.5%	1% (1 item)
Inside a complex coda	2%	–

- 19 This, again, can be read as a preference for a type of splitting that preserves syllabic constituency in order to maximize the recognition of the clipped source elements.

## 2.4. Phonological headedness

- 20 English blends tend to have the phonological contour of at least one of their constituents, i.e. to be parasyllabic with and have the same stress pattern as one or more source elements. This is one more structural feature that enhances the formal recognizability of those source elements which do not appear in full in the blended output. The element determining some of the properties of a whole structure may be termed the head element; here, the source element which gives its phonological contour to the blend is thus seen as its phonological head. If the head source element appears word-initially in the blend, as in the nonce blends in (7), the blend is said to be phonologically left-headed; if it appears in a word-final position, as in (8), it is right-headed:

(7a) *abracadabra* + *caramba* > *abracaramba*

(7b) *avatar* + *turd* > *avaturd*

(8a) *smock* + *apron* > *smapron*

(8b) *stress* + *Cinderella* > *Stresserella*

- 21 If the blend has the same phonological contour as its two source elements, it is said to display ambiheadedness, as illustrated in (9):

(9) *tomato* + *tobacco* > *tomacco*

- 22 The distribution of the various patterns of phonological headedness is presented in Table 4 and contrasted with that available for institutionalized blends (Renner [2019: 42]).

**Table 4: Distribution of English blends according to the type of phonological headedness**

Type of phonological headedness	Institutionalized blends	Nonce blends
left-headedness	24%	14.5% (34 items)
right-headedness	55%	64.5% (150 items)
ambiheadedness	14.5%	4.5% (10 items)
non-headedness	6.5%	16.5% (39 items)

- 23 An overwhelming majority of nonce blends are headed – 83.5% – but the proportion of phonologically non-headed blends is markedly higher in nonce blends than in institutionalized blends. This indicates that in a relatively larger number of blending operations, maximizing the size of the two fragments in the nonce blend has been preferred to replicating the contour of either source element, as illustrated in (10):

(10a) *Dakota* + *Oklahoma* > *Dakotalahoma* (rather than e.g. \**Daklahoma*)

(10b) *Cuba* + *orgasm* > *Cubagasm* (rather than e.g. \**Cubasm*)

- 24 This illustrates a case of competition between two antagonistic recognizability factors and size maximization might be claimed to dominate contour homology more frequently in nonce blending because it is a stronger agent of recognizability



(phonological matter seems to be considered to be more helpful than phonological structure in the correct identification of source elements).

- 25 The two distributions of headed blends are similar as far as the ranking of the three types is concerned, but the gaps between types are wider in nonce blends, with an even more overwhelming preference for right-headedness. One possible partial explanation for the general preference for right-headedness over left-headedness is that it favors an alignment of the stressed syllables of the two source elements word-initially (a majority of English plurisyllabic words are stressed word-initially; this is true of an overwhelming majority of disyllabic words and of a significant majority of trisyllabic words, i.e. the two largest classes of plurisyllabic words [Clopper 2002]).
- 26 Finally, it is of note that a special affinity between the type of lexical shortening and the type of phonological headedness appears in the nonce-blend data, as shown in Table 5.

**Table 5: Distribution of English nonce blends according to the type of shortening and phonological headedness**

Number of nonce blends	left- headedness	right- headedness	ambi- headedness	non- headedness
left-hand-side inner shortening	27	1	2	10
right-hand-side inner shortening	–	111	–	18
double inner shortening	1	13	8	4
haplologic blending	–	22	–	7

- 27 Nonce blends coined through left-hand-side inner shortening are preferentially left-headed, as illustrated in (11), while those coined through right-hand-side inner shortening and haplology are massively right-headed, as illustrated in, respectively, (12) and (13):

- (11a) *casserole* + *loaf* > *casseloaf*  
 (11b) *Nostradamus* + *dumbass* > *Nostradumbass*  
 (12a) *meat* + *catapult* > *meatapult*  
 (12b) *black* + *Frankenstein* > *Blackenstein*  
 (13a) *clam* + *amphitheater* > *Clamphitheater*  
 (13b) *Kent* + *entertainment* > *Kentertainment*

- 28 This distribution is highly significant. In the case of left- and right-hand-side inner shortening, it shows a marked preference for a balance between segmental preservation (for the shorter source element) and contour preservation (for the longer source element) which optimizes the recognition of both source elements and the compactness of the output of blending. In the case of haplologic blending, the longer source element frequently begins with a vowel in our data, as in (13) above, and the act

of blending is an ingenious form of consonantal prothesis which, again, optimizes recognition and compactness.

### 3. Rendering English blends in French

#### 3.1. Formal considerations

##### 3.1.1. Translational typology

- 29 While there is no requirement that a blend in one language be translated as a blend in another, we decided to examine whether the 217 English blends for which a translation was available were rendered as blends in French. A total of 119 corresponding French blends were identified. We adopted a formal and coiner-oriented (i.e. translator-oriented) perspective when classifying them and consequently included several outputs where one source element was English. These bilingual blends were retained even though it is uncertain whether non-bilingual audiences would be able to identify both source elements and recognize the presence of a blend in the examples in (14):

(14a) *clamphithéâtre* < *clam* + *amphithéâtre*

(14b) *ribwich* < *rib* + *sandwich*

- 30 French translators also resorted to using a number of other morphological or non-morphological processes of lexical construction. The variety and distribution of formal subtypes is presented in Table 6. All the blends are two-element units, except for *myphonies*, which is not analyzed further in this section because of its marginal status.

Table 6: Distribution of the formal types of translation of the 217 English blends into French

Formal type	Formal subtype, with an illustrative example	Number	Percentage
Blend	Three-element blend Eng./Fr. <i>myphonies</i> < <i>My</i> + <i>iPhone</i> + <i>phoney</i>	1	< 1%
	Copied blend Eng. <i>fruitopia</i> < <i>fruit</i> + <i>utopia</i> => Fr. <i>fruitopie</i> < <i>fruit</i> + <i>utopie</i>	59	27%
	Semi-creative blend Eng. <i>didgeridon't</i> < <i>didgeridoo</i> + <i>don't</i> => Fr. <i>didgeridiot</i> < <i>didgeridoo</i> + <i>idiot</i>	45	21%
	Creative blend Eng. <i>craptacular</i> < <i>crap</i> + <i>spectacular</i> => Fr. <i>merdeilleux</i> < <i>merde</i> 'shit' + <i>merveilleux</i> 'marvellous'	14	6%
Non-blend	Eng. <i>Wheelchairnocchio</i> < <i>wheelchair</i> + <i>Pinocchio</i> => Fr. <i>Pinocchio en fauteuil roulant</i> 'Pinocchio in a wheelchair'	98	45%

- 31 The French blends can be divided into three main groups. The first group is made of 59 units in which both French source elements are formal analogues of the English source elements. Due to their nature as calques, we classified them as copied blends. The second group, consisting of 45 items, contains blends where one French source element is analogous to an English source element, while the other source element is not. These blends were classified as being semi-creative as a creative translation was necessary to render the non-copied source element. The final group of blends consists of 14 blends where each of the French source elements is markedly different from each of the English source elements and the blends were thus classified as creative.
- 32 We also identified several cases where English blends were not translated as blends in French even though the use of analogous source elements in French would seem to pose no problem, as the examples in (15) show:

(15a) Eng. *Homerific* (< *Homer* + *terrific*)

=> Fr. *°Homerifique* (< *Homer* + *terrifique*)

(15b) Eng. *perfection* (< *par* + *perfection*)

=> Fr. *°perfection* (< *par* + *perfection*)

(15c) Eng. *ovulicious* (< *ovule* + *delicious*)

=> Fr. *°ovulicieux* (< *ovule* + *délicieux*)

- 33 This underlines that rendering an English blend by a French blend remained a free choice made by the translators and that a factor like the possibility of coining a highly felicitous blend – with some segmental overlap – was not necessarily a decisive factor.

### 3.1.2. Lexical shortening

- 34 Due to their nature as copies of the English blends, the copied French blends necessarily involve the same shortening patterns in both languages and, for this reason, we shall focus only on the non-copied blends. Although the type of shortening in the semi-creative blends is influenced by the fact that they share one source element with their corresponding English blend and thus are susceptible to following the same shortening pattern, as in (16a-b), this is not necessarily the case, as can be seen in (16c):

(16a) Eng. *fartzilla* (< *fart* + [*God*]zilla)

=> Fr. *proutzilla* (< *prout* 'fart' + [*God*]zilla)

(16b) Eng. *Nerdstrom* (< *nerd* + [*Berg*]strom)

=> Fr. *beurkstrom* (< *beurk* 'yuck' + [*Berg*]strom)

(16c) Eng. *hip-hopsicle* (< *hip-hop* + [*p*]opsicle)

=> Fr. *esquimhop* (< *esquim[au]* 'eskimo ice cream' + [*hip-h*]op)

- 35 When considered from the point of view of the number of shortenings alone, a distributional discrepancy is visible in Table 7: only a small minority of nonce blends – 14% – displays double shortening, compared to the 43% Renner [2019: 33] noted for institutionalized items.

Table 7: Distribution of French blends according to the type of shortening

Type of shortening	Non-copied nonce blends	Institutionalized blends
Single shortening	39 (66%)	50 (53%)

Double shortening	8 (14%)	40 (43%)
Haplologic	9 (15%)	4 (4%)
Sandwich	3 (5%)	–

- 36 This dispreference for double shortening indicates a preference for maintaining at least one source element entirely intact.

### 3.1.3. Phonological split points

- 37 We only analyzed the split points of the non-copied blends, as these, by definition, contain at least one source element unrelated to an English source element. To avoid the ambiguous analyses of blends containing overlapping segments, we determined the point of phonological splitting in the source elements of the 26 semi-creative blends and 4 creative blends where there was no overlap, giving a total of 60 splits. Splitting was determined as occurring in one of the five possible locations: at syllable boundaries (17a), between onset and nucleus (17b), between nucleus and coda (17c), in a complex onset (17d) or in a complex coda (17e):

- (17a) *ca.ca*.*[to.ès]* ‘cockatoo’ + *boy* > *ca.ca.boy*  
 (17b) *es.qui.m[au]* ‘eskimo ice cream’ + *[hip.h]op* > *es.qui.mop*  
 (17c) *Flan.der[s]* + *[e]x.au.cées*<sup>4</sup> ‘granted’ > *Flan.der.xau.cées*  
 (17d) *fan.f[re.luche]* + *[ba.b]iole* > *fanfirole* (Renner [2019: 35])<sup>5</sup>  
 (17e) *Flan.der[s]* + *[e]x.au.cées* ‘granted’ > *Flan.der.xau.cées*

- 38 Table 8 groups the semi-creative and creative blends together as no significant differences were observed between the two categories.

Table 8: Distribution of phonological split points in French non-overlapping blends

Location of split point	Non-copied nonce blends	Institutionalized blends
Syllable boundary	88% (51 items)	47%
Onset-nucleus	9% (5 items)	34%
Nucleus-coda	2% (1 item)	14%
Inside a complex onset	–	6%
Inside a complex coda	2% (1 item)	–

- 39 The distribution demonstrates that nonce blends show a greater preference for splitting to take place at syllable boundaries rather than within syllables as this facilitates the identification of source elements and thus the comprehension of the blend. This distribution also contrasts markedly with that observed by Renner [2019: 35] for institutionalized blends, with the nonce blends showing a strong dispreference for splitting between onset and nucleus.

### 3.1.4. Phonological headedness

- 40 In French, phonological headedness is determined by the relationship between the length of the blend and that of its source elements, length being defined here as the number of syllables. The distribution of the data presented in Table 9 shows that the non-copied nonce blends are fairly similar to the institutionalized blends examined by Renner [2019: 43].

Table 9: Distribution of French blends according to the type of phonological headedness

Type of headedness	Non-copied nonce blends	Institutionalized blends
left-headedness	20% (12 items)	18%
right-headedness	37% (22 items)	25%
ambiheadedness	8% (5 items)	9%
non-headedness	34% (20 items)	48%

- 41 Given that non-headedness means that a blend has the length of neither of its source elements, the gap between institutionalized blends and nonce blends might indicate that the latter tend to increase in length in order to retain as many phonemes from their source elements as possible. To explore this hypothesis, we examined the length of the non-copied blends relative to their longer source element.

Table 10: Distribution of the length of French non-copied blends relative to the length of their longer source element

Relative length	Non-copied nonce blends
Greater than the longer source element	32% (19 items)
Parasyllabic with the longer source element	64% (38 items)
Lesser than the longer source element	3% (2 items)

- 42 The data in Table 10 show a strong dispreference for reducing the number of syllables in a blend, with the marked overall preference being the preservation of the length of the longer source element, as in (18):

- (18a) Eng. *bagzooka* < *bag* + *bazooka*  
 => Fr. *saczooka* < *sac* 'bag' + *bazooka*  
 (18b) Eng. *Oklasoft* < *Oklahoma* + *Microsoft*  
 => Fr. *oclafoutis* < *Oklahoma* + *clafoutis*

- 43 Retaining the length of the longer source element is a factor of increased recognizability, and especially so in the case of French, which, unlike English, cannot use the stress pattern cue in the correct identification of source elements.

### 3.2. Semantic considerations

- 44 To conclude this analysis of the translations of the original Simpsonian blends into French, we also sought to better appreciate the role of semantic considerations in the choice of source elements in the resultant French blends. Accordingly, the copied blends were not considered here, as they did not involve the translators making any free choice in the selection of the French source elements. Additionally, the analogous elements of the semi-creative blends – such as *Godzilla* in (16a) – were also not considered for the same reason. The non-analogous source elements of the French semi-creative and creative blends were divided into three groups. The first group contains blends for which the corresponding English and French source elements are synonyms, like *bag* and *sac* in (18a). The second group contains blends where the source elements are cohyponyms, like *popsicle* and *esquimau* in (16c), which are both frozen dessert foods. The final group contains the remaining items, where the corresponding source elements are not remarkably related semantically, such as *nerd* and *beurk* in (16b). A total of 73 non-analogous source elements for each language were examined. They comprised both source elements from the 14 French creative blends and their English equivalents (28 pairs), as well as the non-analogous source elements from the 45 French semi-creative blends and their English equivalents (45 pairs). Of the 73 pairs of source elements, 22 displayed a relation of synonymy and a further 22 pairs displayed a relation of cohyponymy, as shown in Table 11.

**Table 11: Distribution of the semantic relations between English and French source elements of non-copied blends relative to the presence of overlapping segments in French**

English and French source elements	Synonymous	Cohyponymic	Not remarkably related
Overlapping segments	12	22	29
No overlapping segments	10	–	–

- 45 Table 11 also indicates whether these pairs of source elements allowed segmental overlap of the French elements because of a remarkable cross-distribution of formal and semantic types. In the cases where a French blend can be made through the use of synonyms of the English source elements, there is no clear preference between choosing source elements that allow segmental overlap, as in (19a), and source elements that do not, as in (19b):

- (19a) Eng. *Spellympics* (< *spelling* + *Olympics*)  
 => Fr. *ortholympiques* (< *orthographe* + *olympiques*)  
 (19b) Eng. *Merry Fishmas* (< *Merry Christmas* + *fish*)  
 => Fr. *Poissex Noël* (< *poisson* ‘fish’ + *Joyeux Noël* ‘Merry Christmas’)

- 46 However, in cases where a French blend contains a source element which is not a synonym of the English source element, this element always allows for segmental overlap, as illustrated in (20):

- (20a) Eng. *pray-per-view* (< *pray* + *pay per view*)  
 => Fr. *paradiabolique* (< *para* + *diabolique* + *bolique*)  
 (20b) Eng. *purrplexing* (< *purr* + *perplexing*)  
 => Fr. *chat roule* < *chat* 'cat' + *ça roule* 'cool'

- 47 As a whole, the data in Table 11 underline the importance of overlapping syllables, which help with the identification of the source elements of nonce blends. As to the striking constraint of overlapping for the translated French units not made of source elements which are synonyms of the original English source elements, the explanation might be that translators either choose to primarily respect the meaning of each original source element, in which case formal felicitousness (typified by overlapping) is secondary, or to primarily favour the formal felicitousness of the output, in which case a semantically precise translation (through the use of synonyms) is not essential.

## Conclusion

- 48 The examination of 237 English Simpsonian blends demonstrated several formal differences between nonce blends and institutionalized blends, notably in terms of the preferred type of lexical shortening and the prevalence of overlapping segments and phonological headedness. These particularities of nonce blends combine to increase the recognizability of the source elements and thus enhance the understanding of the novel output. The English nonce blends were rendered as French blends in slightly over one case in two. The latter showed highly similar tendencies to the English blends, particularly in the choice of split points, shortening patterns and phonological headedness. These results indicate that, against a widely-held view among morphologists (cf. 2.1. above), blends may not constitute a homogeneous class from a formal standpoint and that additional fine-grained studies investigating the intricacies of blending in English and other languages are undoubtedly still in order.

---

## BIBLIOGRAPHY

- BAUER Laurie, 2012, "Blends: Core and periphery", in RENNER Vincent, MANIEZ François & ARNAUD Pierre (Eds.), *Cross-disciplinary Perspectives on Lexical Blending*, Berlin: De Gruyter Mouton, 11-22.
- BELIAEVA Natalia, 2014, *Unpacking Contemporary English Blends: Morphological Structure, Meaning, Processing*, PhD dissertation, Victoria University of Wellington.
- BELIAEVA Natalia, 2019, "Blending in morphology", in ARONOFF Mark (Ed.-in-chief), *Oxford Research Encyclopedia of Linguistics*, Oxford: Oxford University Press, available at <https://oxfordre.com/linguistics/abstract/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-511>
- CLOPPER Cindy, 2002, "Frequency of stress patterns in English: A computational analysis", *IULC Working Papers*, 2 (1), Bloomington: Indiana University.

HOHENHAUS Peter, 2005, "Lexicalization and institutionalization", in ŠTEKAUER Pavol & LIEBER Rochelle (Eds.), *Handbook of Word-Formation*, Dordrecht: Springer, 353-373.

LÓPEZ-RÚA Paula, 2004, "The categorial continuum of English blends", *English Studies*, 85 (1), Abingdon: Routledge, 63-76.

OED = *Oxford English Dictionary Online*, available at <https://www.oed.com>

RENNER Vincent, 2015, « Panorama rétro-prospectif des études amalgamatives », *Neologica*, 9, Paris: Classiques Garnier, 97-112.

RENNER Vincent, 2019, "French and English lexical blends in contrast", *Languages in Contrast*, 19 (1), Amsterdam: Benjamins, 27-47.

RONNEBERGER-SIBOLD Elke, 2006, "Lexical blends: Functionally tuning the transparency of complex words", *Folia Linguistica*, 40 (1-2), Berlin: Mouton de Gruyter, 155-181.

## NOTES

1. The original coinage of this term has likely been captured in the earliest attestation quoted in the OED (*Transactions of the Electrochemical Society* [1929]): "Because of the frequent use of the terms 'adsorbed ion' and 'adsorbed atom' we would suggest that they be abbreviated to adion and adatom".
2. Periods indicate syllable boundaries and the relevant syllable is underlined in case of double shortening.
3. No example of splitting at this point is attested in the dataset of English Simpsonian blends.
4. Or, more accurately, in phonological terms: [ɛ]g.zo.se.
5. No example of splitting at this point is attested in the dataset of French translations.

## ABSTRACTS

This contribution examines the conspicuous presence of lexical blends in the long-running US television show *The Simpsons* and consists of two parts. The first part involves the formal analysis of 237 nonce blends in the original English-language version of the show, working on the underlying hypothesis that, despite their novelty, the audience is nonetheless able to easily decipher the blends due to a number of formal choices enhancing the recognizability of their source elements. The second part then examines the translation of these blends into Hexagonal French by taking account of formal and semantic considerations influencing whether the English nonce blends are rendered as blends in French and, if so, whether the latter display the same formal tendencies as in English. It is found that Simpsonian nonce blends notably stand out in terms of preferred type of lexical shortening and prevalence of segment overlap and phonological headedness, in both English and French. These results indicate that, against a widely-held view among morphologists, blends may not constitute a homogeneous class from a formal standpoint.



Cet article s'intéresse aux amalgames lexicaux de la série télévisée américaine *Les Simpson* et est divisé en deux parties. La première partie propose une analyse formelle des 237 occasionnalismes relevés dans la version originale anglaise de la série et se fonde sur l'hypothèse que, bien que ces amalgames soient de nouveaux mots construits inconnus, ils sont compris par les téléspectateurs du fait de plusieurs choix structurels facilitant la reconnaissance des différents éléments-sources. La deuxième partie de l'article examine ensuite les choix de traduction des amalgames anglais dans la version française de la série. Elle s'intéresse plus particulièrement à différents facteurs formels et sémantiques influençant ces choix et au degré de similitude entre les amalgames de l'anglais et du français. Il est conclu que les occasionnalismes amalgamés simpsoniens se distinguent tendanciellement des amalgames institutionnalisés, notamment pour ce qui concerne la distribution des patrons d'accourcissement, les choix de chevauchement segmental entre éléments-sources et les préférences en termes de tête phonologique. Ces résultats indiquent que, contrairement à un point de vue dominant chez les morphologues, les amalgames ne semblent pas former une classe homogène du point de vue formel.

## INDEX

**Keywords:** word-formation, lexical blending, nonce word, English, French

**Mots-clés:** morphologie, amalgame lexical, mot-valise, occasionnalisme, anglais, français

## AUTHORS

### ADAM RENWICK

University of Lyon (Lumière Lyon 2)  
adam.renwick1@univ-lyon2.fr

### VINCENT RENNER

University of Lyon (Lumière Lyon 2)  
vincent.renner@univ-lyon2.fr